Reinforcement or Revenue? Exploring the Impact of Robot Response Style on Customer Trust, Satisfaction, and Tipping

Fahad Dalwai
Department of Computer Science
University of Illinois Chicago
Chicago, Illinois, United States
fdalw@uic.edu

Farhan Saif
Department of Computer Science
University of Illinois Chicago
Chicago, Illinois, United States
fsaif@uic.edu

Abstract—Service robots are increasingly being introduced in restaurants, yet little is known about how their response style influences customer experience and behavior. This study employs a between-subjects design with a Misty II robot comparing two scripted strategies during order placement: positive reinforcement (affirming choices) versus upselling (suggesting add-ons). Participants followed a simple restaurant-style flow by placing an order from a paper menu, then completing a brief post-interaction survey and selecting a tip amount on a mock bill. Through an exploratory mixed-methods approach, this study investigates how socially supportive vs commercially persuasive dialogue affects perceptions and spending.

Index Terms—Human-Robot Interaction, Service Robots, Customer Satisfaction, Trust, Tipping Behavior, Positive Reinforcement, Upselling

I. INTRODUCTION

Over the past two decades, the presence of robots in customer-facing service roles—particularly in restaurants, bars, and hotels—has grown rapidly as businesses seek to improve efficiency and consistency [1], [2]. Yet effective service is not just about speed or accuracy; it also depends on how communication feels to the customer. Studies of human servers show that how you talk matters: writing "thank you" on checks increases tips [3], and better service leads to bigger tips [4]. However, attempts to upsell can succeed or backfire depending on how customers interpret the intent behind them. Whether robots elicit similar reactions remains largely unknown.

Most current robot research focus on aspects such as politeness and emotion [5], while few examine sales-oriented communication. When a robot says "Great choice!" versus "Would you like fries with that?", does it change how customers feel—or how much they tip? Compliments may enhance perceived warmth and trust, whereas overt sales pitches could reduce likability or perceived sincerity. Understanding these effects is crucial for designing robots that balance positive user experience with business outcomes.

We compare two common response styles on Misty II during order placement: positive reinforcement vs upselling—in a realistic café-style interaction. Stated formally, our paper asks the following novel research question as such;

> *'Do compliments or sales pitches from a robot change how satisfied customers are, how much they trust it, and how much they tip?'*

This exploratory study uses a realistic restaurant scenario to extend findings from human service research to HRI, providing early insights into how conversational framing shapes both social perception and economic behavior in robotic service contexts.

II. RELATED WORK

This study builds on four areas: service robots in restaurants, positive reinforcement, upselling strategies, and trust and tipping. Together, these show how different ways of talking might affect satisfaction and tips.

A. Service Robots in Restaurants

Service robots are now being used in restaurants and retail to make things more consistent. Studies of restaurant robots like Pepper and Relay show that being polite, showing emotion, and talking like a human make customers more willing to interact [5]. But most research looks at building relationships—like warmth or empathy—not the sales talk that servers actually use in real restaurants.

B. Positive Reinforcement with Robots and Human Servers

Positive reinforcement means praising people when they do something, which makes them want to do it again. In robot research, warmth and competence have been shown to increase trust [6]. For human servers, the evidence is clear: waiters who complimented customers ("You made an excellent choice") got higher tips and better ratings [3], [7]. This kind of talk makes people feel good and improves how they see the service. But we don't know if it works the same way when a robot does it.

C. Upselling Strategies

Upselling means trying to get customers to buy more stuff or upgrades. With human servers, whether it works depends on how customers see it: if it feels helpful, they're more likely to say yes, but if it feels pushy, they get annoyed.

D. Trust and Tipping

Trust matters a lot for both satisfaction and whether people do what the robot suggests [6]. In restaurants, trust directly affects how customers see the service and whether they come back. Tipping gives us a real behavior we can measure. Research on human servers shows that friendliness and gratitude increase tips [4]. Related work on robotic bartenders found that only about 11% of participants said they would tip a robot, often reasoning that robots do not need income or appreciation; however, willingness to tip increased when participants learned that tips would benefit human staff or help the robot improve its service [10]. By measuring both tips and survey responses, this study looks at both the economic and psychological sides of robot-customer interaction.

E. Summary

Past research shows that expression of talk, warmth, and trust all matter for satisfaction with both human and robot service. But no study has directly compared positive reinforcement and upselling when a robot takes orders. Building on what we know about human servers, this study looks at how these two ways of talking affect both how customers feel and how much they tip.

III. DESIGN FACTORS

The goal of this study is to examine how a restaurant robot's response strategy influences customer experience. As previously mentioned, two contrasting interaction styles—positive reinforcement and upselling—were selected to represent socially supportive versus commercially persuasive communication. These strategies are conceptually distinct yet valid within service contexts: positive reinforcement aligns with social rapport-building behaviors shown to enhance satisfaction and trust [3], [6], [11], while upselling reflects persuasive sales techniques commonly used in hospitality to increase revenue [12], [13]. The following design factors guided the study.

A. Response Strategy

A robot's verbal response strategy can shape how users interpret its role: Positive reinforcement involves the robot affirming the customer's choice (e.g., "Great choice! That's one of my favorites."). This form of feedback conveys social warmth and appreciation, behaviors that have been linked to increased satisfaction and perceived competence in prior HRI studies [14], [15].

Upselling, conversely, involves suggesting additional items related to the user's choice (e.g., "Would you like to add fries or a drink with that?"). While common in human service interactions [16], upselling may be perceived as instrumental or intrusive when performed by a robot [17].

By contrasting these two strategies, the study aims to identify how social affirmation versus commercial persuasion influences customers' subjective and behavioral responses.

B. Perception Dimensions

Previous work suggests that human responses to service robots are shaped by multi-dimensional perceptions including trust, acceptance, and proximity.

Trust in human-robot interaction is conceptualized as a multifaceted construct emerging from dynamic user-robot relationships. Kok and Soh [18] define trust as essential for successful robot deployment, noting that insufficient trust leads to underutilization while excessive trust raises safety concerns due to inadequate monitoring. Hancock et al.'s [6] framework, widely adopted in both reviews, identifies three factor categories: human-related (personality traits, expertise), robot-related (behavior, reliability, anthropomorphism), and environmental (task characteristics, communication). Measurement approaches include standardized questionnaires like the Trust Perception Scale-HRI [20], behavioral indicators such as task intervention and advice-following, and emerging sensor-based methods utilizing physiological measurements and machine learning for real-time trust assessment during ongoing interactions.

Babel et al. [21] investigated how a social robot's dialog content, talk initiative, and gaze behavior influence trust, acceptance, and proximity in human-robot interaction, finding that proactive communication strategies significantly enhance trust and acceptance across various conversation themes, including service tasks. For restaurant robots taking orders and providing suggestions, this research indicates that proactive behavior—such as initiating menu recommendations or asking about dietary preferences—would increase customer acceptance and satisfaction more than passive order-taking. The study showed that "different gaze and proactive strategies enhanced trust and acceptance across various dialog content," suggesting restaurant robots combining proactive suggestion-making with appropriate gaze behavior during food-related conversations foster higher customer acceptance and more natural, service-oriented interactions.

These dimensions collectively describe the user's experience quality and are operationalized through validated survey measures after interaction.

C. Economic Behavior

Tipping behavior in service robot settings provides a concrete behavioral measure that adds to subjective trust and satisfaction assessments. Campagna and Rehm's systematic review identifies behavioral trust measurement approaches including "behavioral change" where "certain physical cues exhibited by the human during the interaction can give insights into their level of trust," with tipping showing customers' financial expression of trust and satisfaction beyond verbal feedback. The review highlights that "trust manifests through observable behaviors" including following advice, task delegation, and decision-making under uncertainty, with tipping matching

these indicators as customers choose whether to reward robot service. Measurement-wise, recording tipping requires tracking both occurrence (yes/no) and amount (dollar value), then linking these with subjective measures from proven tools like Schaefer's Trust Perception Scale-HRI. The systematic review notes that "multi-modal behavioral measures analyze trust through body postures, facial expressions, and movement patterns," suggesting tipping data could be combined with visible engagement behaviors during service. However, robot human-likeness creates distinct challenges, as studies by Natarajan and Gombolay [22] on anthropomorphism effects, and Parenti et al. [23] on human-like feedback gestures, show these features shape trust perceptions and decision processes. If customers don't see robots as needing tips due to lacking financial needs, tipping rates may stay uniformly low regardless of service quality, creating measurement limits.

IV. Apparatus & Setup

The study will be conducted using a Misty II robot positioned on a tabletop to emulate a restaurant server stationed beside the customer. Misty II is a socially expressive platform equipped with a front-facing display for animated eyes, an onboard speaker, and a microphone array, enabling speech synthesis and two-way verbal interaction. Its coordinated head movements and subtle facial animations were used to convey attentiveness and acknowledgment, supporting a natural flow of conversation during the ordering task. Prior work has shown that a robot's communication style and nonverbal expressiveness significantly influence user engagement and comfort [24], which makes Misty an appropriate choice for testing socially oriented interaction styles in this experiment.

A. Hardware & Placement

The experiment would be conducted in a quiet laboratory configured to resemble a small café table setting. Each participant was seated at a table with a paper menu displaying a fixed list of six food and beverage items with corresponding prices. The Misty II robot was positioned across from the participant at approximately eye level to support natural conversational engagement [25], [26]. Soft ambient café sounds played continuously to enhance ecological realism.

A concealed experimenter station controlled Misty's dialogue flow from a laptop connected via Wi-Fi. Participants were unaware of both the condition assignment and the control interface. All robot responses were generated live using the GPT-4 model via an API connection to a custom Misty control interface. Prompts were condition-specific to ensure consistent tone and structure across trials, and the generated text was immediately vocalized through Misty's built-in speech synthesis module.

B. Technical Implementation

Misty's dialogue will be powered by MistyGPT, a lightweight backend built on OpenAI's GPT-based API for structured conversational control. The robot's speech output will be synchronized with subtle head nods and facial expression changes (smiling eyes) to mimic human server behavior. Interaction logs will also be automatically recorded, capturing timestamps and session duration, text based interactions.

V. Measures & Analysis

This study employed both quantitative and qualitative measures to assess how the robot's response strategy shaped customer experience. Data included self-report ratings, behavioral tipping outcomes, and open-ended comments collected immediately after each interaction.

A. Quantitative Measures

Customer Satisfaction. Participants will rate their overall satisfaction with the interaction using three 5-point Likert items adapted from prior HRI satisfaction studies [6], [7], [27]. Items assessed enjoyment ("I enjoyed interacting with Misty"), fulfillment ("The interaction met my expectations"), and perceived quality ("The service felt satisfying").

Trust in the Robot. Trust was measured using the Human–Machine Trust Scale developed by Jian et al. (2000) [28], widely adopted in HRI research [29], [30]. Seven items captured dependability, honesty, and perceived reliability (e.g., "I can rely on the robot to act in my best interest").

Perceived Friendliness and Usability. Friendliness was rated using two 5-point items (e.g., "The robot seemed friendly," "The robot was pleasant to talk to"). Usability was measured with three items adapted from the System Usability Scale [31], reworded for short conversational tasks (e.g., "It was easy to communicate with the robot").

Tipping Behavior. Participants indicated a tip amount on a printed mock payment slip (options: 0%, 10%, 15%, 20%, or custom). The recorded amount served as a behavioral measure of economic engagement, complementing self-reported evaluations and allowing comparison of social versus commercial influence across response conditions.

B. Qualitative Measures

Open-ended questions asked participants to share their impressions of the interaction (e.g., "What did you think about the robot's responses?"). These reflections help capture emotional tone, perceived intent, and reactions to the robot's communication style. Responses will be examined through thematic analysis [32] to identify recurring patterns, complementing the quantitative results and providing deeper insight into how participants interpreted the robot's behavior.

C. Analysis Plan

Quantitative data will be analyzed using independent-samples t-tests to compare mean differences between the Positive Reinforcement and Upselling conditions across all dependent variables (satisfaction, trust, friendliness, usability, and tipping). This test is appropriate because participants were randomly assigned to one of two independent groups [30]. When assumptions of normality or homogeneity of variance are violated, the non-parametric Mann–Whitney U test will be used as a robust alternative [31].

To explore relationships among experience dimensions, Pearson correlation analyses (or Spearman's ρ when data are non-normal) will be conducted between satisfaction, trust, and tipping behavior.

Qualitative comments will be analyzed through inductive thematic coding conducted independently by two researchers to ensure objectivity. Inter-rater reliability will be assessed using Cohen's κ. Emergent themes will then be integrated with quantitative patterns to triangulate findings, helping us explain why certain interaction styles influenced participants' perceptions and behaviors. All the participant resonses will be taken using Qualtrics due to its ease of usage.

D. Unit of Analysis

Each participant's single interaction with Misty constituted one dyadic interaction unit. Using a single-participant (dyadic) setup avoids social confounds from group dynamics, as prior research has shown that larger group sizes are often associated with higher tip amounts [33]. To avoid this confounding effect, this study focuses on one-on-one customer–robot interactions as the primary unit of analysis, with no repeated exposures across conditions.

VI. PARTICIPANTS & PROCEDURE

A. Participants

Participants were recruited from the university campus. Ideally, the study aims to include approximately ten adult participants to allow for preliminary comparison between conditions. All participants will report fluent English proficiency and no prior experience interacting with the Misty robot. Participants will be informed that the study involves evaluating a robot ordering system for a café prototype, without mention of upselling or reinforcement strategies to minimize expectancy bias [32].

B. Procedure

Participants completed a short pre-interaction demographic survey before beginning the session. Each session followed a structured sequence lasting approximately 10–12 minutes from arrival to survey completion.

Arrival and Orientation: Participants arrived individually and were welcomed by the experimenter. They were seated at a café-style table facing the Misty II robot. Background ambient audio (light chatter and soft café music) played to simulate a casual restaurant atmosphere.

Consent and Cover Story: Participants were told that the study examined robot ordering interactions. They received a brief overview of the ordering process and were asked to imagine being customers at a small café testing a new robotic server.

Condition Assignment: Each participant was randomly assigned to one of two between-subject conditions: Positive Reinforcement or Upselling. Assignment was concealed, and participants were unaware of the differing response types.

Interaction Phase: Misty initiated the conversation with a standardized greeting ("Hello! Welcome to the café. What would you like to order today?"). In the Positive Reinforcement condition, Misty responded with affirming statements matched to the participant's choices (e.g., "Excellent choice!" or "That's one of our popular items."). In the Upselling condition, Misty made additional item suggestions related to the participant's order (e.g., "Would you like to add a drink with that?" or "Would you like a dessert to go with your meal?"). Misty continued the dialogue in each condition until the participant finalized their order and confirmed that they were ready to proceed. Both conditions were designed to maintain consistent response tone, pacing, and overall length. Participants could ask short clarification questions if needed.

Order Confirmation and Payment: Misty confirmed the order, read the total, and thanked the participant. A mock payment slip was then placed on the table, displaying a bill and optional tip percentage selections. Participants circled or wrote a tip amount.

Post-Interaction Survey: After completing payment, participants filled out a brief paper questionnaire measuring satisfaction, trust, friendliness, and usability, followed by two open-ended feedback questions about the interaction.

C. Randomization and Control

Participants will be randomly assigned to conditions using a counterbalanced assignment sequence. The experimenter controlling Misty will be blind to participant identity and followed a fixed dialogue script to maintain consistency in tone, timing, and turn-taking across all trials. No spontaneous or adaptive language was used to ensure reproducibility.

Environmental variables such as lighting, background sound, and seating distance will be held constant across sessions. All sessions and interactions will be recorded via Misty's onboard sensors and an external camera for verification of dialogue timing and participant responses, consistent with prior data-logging practices in controlled HRI research [26].

VII. LIMITATIONS

Due to the limitations and lack of practicality, our study without doubt has a few limitations. The small convenience sample of university participants limits statistical power and generalizability beyond this population. The lab café setup improves control but lacks the social and environmental realism of an actual restaurant. Tipping was simulated on a mock bill rather than involving real money, which may weaken ecological validity. Because only the Misty II platform was used, findings may not extend to other robot embodiments or voices. The scripted dialogue ensured consistency but reduced natural conversational adaptation. Finally, interactions occurred in isolation rather than group settings, omitting social influences such as peer presence and group tipping norms. Future work should address these constraints through larger field deployments with real payments and varied robot designs.

REFERENCES

[1] S. Ivanov and C. Webster, "Adoption of robots and service automation by tourism and hospitality companies," *Rev. Turismo & Desenvolvimento*, vol. 27/28, pp. 1501–1517, 2017.

[2] G. Odekerken-Schröder, C. Mele, T. Russo-Spena, *et al.*, "The service triad: An empirical study of service robots, customers and frontline employees in a fast casual dining restaurant," *J. Service Manag.*, 2022.

[3] B. Rind and P. Bordia, "Effect of server's 'thank you' and personalization on restaurant tipping," *J. Appl. Soc. Psychol.*, vol. 25, no. 9, pp. 745–751, 1995, doi: 10.1111/j.1559-1816.1995.tb01772.x.

[4] M. Lynn and M. McCall, "Gratitude and gratuity: A meta-analysis of research on the service–tipping relationship," *J. Socio-Econ.*, vol. 29, pp. 203–214, 2000.

[5] M. Salem, M. Ziadee, and M. Sakr, "Effects of politeness and interaction context on perception and experience of HRI," in *Proc. Int. Conf. Social Robotics (ICSR)*, 2013, pp. 531–541.

[6] P. A. Hancock, D. R. Billings, K. E. Schaefer, *et al.*, "A meta-analysis of factors affecting trust in human–robot interaction," *Human Factors*, vol. 53, no. 5, pp. 517–527, 2011.

[7] A. Parasuraman, V. A. Zeithaml, and L. L. Berry, "SERVQUAL: A multiple-item scale for measuring consumer perceptions of service quality," *J. Retail.*, vol. 64, no. 1, pp. 12–40, 1988.

[8] L. Cain, M. Busser, and M. Kang, "Robo-tipping: Are customers game?," in *Advances in Hospitality and Leisure*, vol. 17, Cham, Switzerland: Springer, 2021, pp. 125–141.

[9] J. S. Seiter and R. L. Gass, *Persuasion, Social Influence, and Compliance Gaining*. Boston, MA, USA: Allyn & Bacon, 2005.

[10] M. McCall and C. Voorhees, "The drivers of loyalty program success: An organizing framework and research agenda," *Cornell Hosp. Q.*, vol. 51, no. 1, pp. 35–52, 2010.

[11] J. J. Ahn, S. L. Kim, and S. J. Lee, "Promoting hotel upselling: The effect of message appeal and modality," *J. Hospitality Tourism Manag.*, vol. 50, pp. 64–73, 2022.

[12] G. Briggs, T. Williams, and M. Scheutz, "Toward socially-sensitive utterance generation in human–robot interaction," in *AAAI Spring Symp. Series*, 2016.

[13] S. M. Fiore, T. J. Wiltshire, E. J. C. Lobato, *et al.*, "Toward understanding social cues and signals in human–robot interaction," *Front. Psychol.*, vol. 4, p. 865, 2013.

[14] J. Ni and J. Jian, "Upselling versus upsetting customers? A model of intrinsic and strategic upselling," Working Paper, Wharton School, Univ. Pennsylvania, 2015.

[15] M. F. Jung, N. Martelaro, and P. J. Hinds, "Affective grounding in human–robot interaction," in *Proc. ACM/IEEE Int. Conf. Human–Robot Interaction (HRI)*, 2015, pp. 263–270.

[16] B. C. Kok and H. Soh, "Trust in robots: Challenges and opportunities," *Curr. Robot. Rep.*, vol. 1, pp. 297–309, 2020.

[17] K. E. Schaefer, "Measuring trust in human–robot interactions: Development of the 'trust perception scale–HRI'," in *Robust Intelligence and Trust in Autonomous Systems*, R. Mittu, D. Sofge, A. Wagner, and W. Lawless, Eds. Cham, Switzerland: Springer, 2016, pp. 191–218.

[18] F. Babel, S. Kraus, M. Miller, *et al.*, "Can we trust a robot like a physician? A field study to understand the social acceptance of pain management robots," in *Proc. ACM/IEEE Int. Conf. Human–Robot Interaction (HRI)*, 2021, pp. 58–66.

[19] M. Natarajan and M. Gombolay, "Effects of anthropomorphism and accountability on trust in human–robot interaction," in *Proc. ACM/IEEE Int. Conf. Human–Robot Interaction (HRI)*, 2020, pp. 33–42.

[20] L. Parenti, A. W. Lukomski, D. De Tommaso, M. Belkaid, and A. Wykowska, "Human-likeness of feedback gestures affects decision processes and subjective trust," *Int. J. Social Robotics*, pp. 1–9, 2022.

[21] S. Saunderson and G. Nejat, "Persuasive robots should avoid authority: The effects of role and communication style," *Sci. Robotics*, vol. 6, no. 52, 2021.

[22] M. Brengman, K. Willems, and H. Van Kerrebroeck, "An observational study comparing a humanoid service robot and a tablet kiosk at POS," *J. Bus. Res.*, vol. 122, pp. 423–430, 2021.

[23] S. Edirisinghe, M. Papadopoulos, and K. Dautenhahn, "Field trial of a shopworker robot with friendly guidance," *ACM Trans. Hum.-Robot Interact.*, 2023.

[24] R. L. Oliver, "A cognitive model of the antecedents and consequences of satisfaction decisions," *J. Mark. Res.*, vol. 17, no. 4, pp. 460–469, 1980.

[25] J.-Y. Jian, A. M. Bisantz, and C. G. Drury, "Foundations for an empirically determined scale of trust in automated systems," *Int. J. Cogn. Ergonom.*, vol. 4, no. 1, pp. 53–71, 2000.

[26] K. E. Schaefer, J. Y. Chen, J. L. Szalma, and P. A. Hancock, "A meta-analysis of factors influencing the development of trust in automation," U.S. Army Research Laboratory Tech. Rep., 2013.

[27] S. C. Kohn, E. A. Parker, J. Y. C. Chen, *et al.*, "Measurement of trust in automation: A narrative review and reference guide," *Human Factors*, vol. 63, no. 7, pp. 1013–1057, 2021.

[28] J. Brooke, "SUS: A 'quick and dirty' usability scale," in *Usability Evaluation in Industry*, P. W. Jordan, B. Thomas, B. A. Weerdmeester, and I. L. McClelland, Eds. London, U.K.: Taylor & Francis, 1996, pp. 189–194.

[29] V. Braun and V. Clarke, "Using thematic analysis in psychology," *Qual. Res. Psychol.*, vol. 3, no. 2, pp. 77–101, 2006.

[30] A. Field, *Discovering Statistics Using IBM SPSS Statistics*, 4th ed. Los Angeles, CA, USA: Sage, 2013.

[31] J. Pallant, *SPSS Survival Manual: A Step by Step Guide to Data Analysis Using IBM SPSS*, 7th ed. New York, NY, USA: Routledge, 2020.

[32] M. Strait, P. Briggs, and M. Scheutz, "Gender, more so than age, modulates positive perceptions of language-based HRI," in *Proc. AISB*, 2015.

[33] N. Jahan, "Determinants of Tipping Behavior: Evidence from US Restaurants," M.S. thesis, Dept. of Economics, South Dakota State Univ., Brookings, SD, USA, 2018.